

## SOMMAIRE

<b>CHOIX DE LA METHODE D'INTERPOLATION SPATIALE</b>	<b>3</b>
<b>CARTOGRAPHIE DES CHAMPS DE CONCENTRATION EN DIOXYDE D'AZOTE (NO<sub>2</sub>) ET EN BENZENE (C<sub>6</sub>H<sub>6</sub>) - INTERPOLATION PAR KRIGEAGE ORDINAIRE MONOVARIABLE</b>	<b>4</b>
Analyse du comportement spatial des données et modélisation statistique	4
Contrôle de la précision du modèle d'interpolation	5
Résultats cartographiques	6
<b>INTEGRATION D'INFORMATIONS INDIRECTES LIEES A DES VARIABLES AUXILIAIRES</b>	<b>9</b>
<b>Cartographie des champs de concentration en dioxyde d'azote (NO<sub>2</sub>) et en benzène (C<sub>6</sub>H<sub>6</sub>) avec prise en compte de données auxiliaires - Interpolation par krigeage multivariable</b>	<b>9</b>
<i>Choix de la méthode d'interpolation en krigeage multivariable</i>	9
Vérification de la précision et de l'optimisation du modèle d'interpolation	11
Résultats cartographiques	12
<b>CONCLUSION</b>	<b>15</b>

L'interpolation spatiale permet à l'aide de fonctions mathématiques d'estimer des valeurs de pollution en tous points ou nœuds d'une grille régulière superposée à un domaine d'étude sur la base de mesures ponctuelles et dispersées. Cette cartographie spatialisée à échelle fine de la qualité de l'air sur une aire géographique donnée répond aux recommandations de la loi sur l'air du 30 décembre 1996 ainsi qu'aux besoins croissants des collectivités locales et régionales. Elle repose sur l'application de méthodes spécifiques de spatialisation des données et l'utilisation de logiciels adaptés.

### Choix de la méthode d'interpolation spatiale

Les algorithmes d'interpolation spatiale couramment utilisés jusqu'à récemment en matière de cartographie des champs de pollution concernaient pour l'essentiel des méthodes de type déterministe comme par exemple les méthodes de triangulation, de Shepard, de splines laplaciennes ou d'inverse des distances.

Chacun de ces algorithmes utilise des techniques différentes de pondération entre les variables à cartographier et aboutissent de ce fait à des résultats différents. En règle générale, les résultats de ces algorithmes d'interpolation sont issus d'une règle de pondération calculée sur la base des concentrations observées aux sites de mesure voisins du nœud le plus proche. Ils ne prennent pas en compte les caractéristiques spatiales naturelles du polluant.

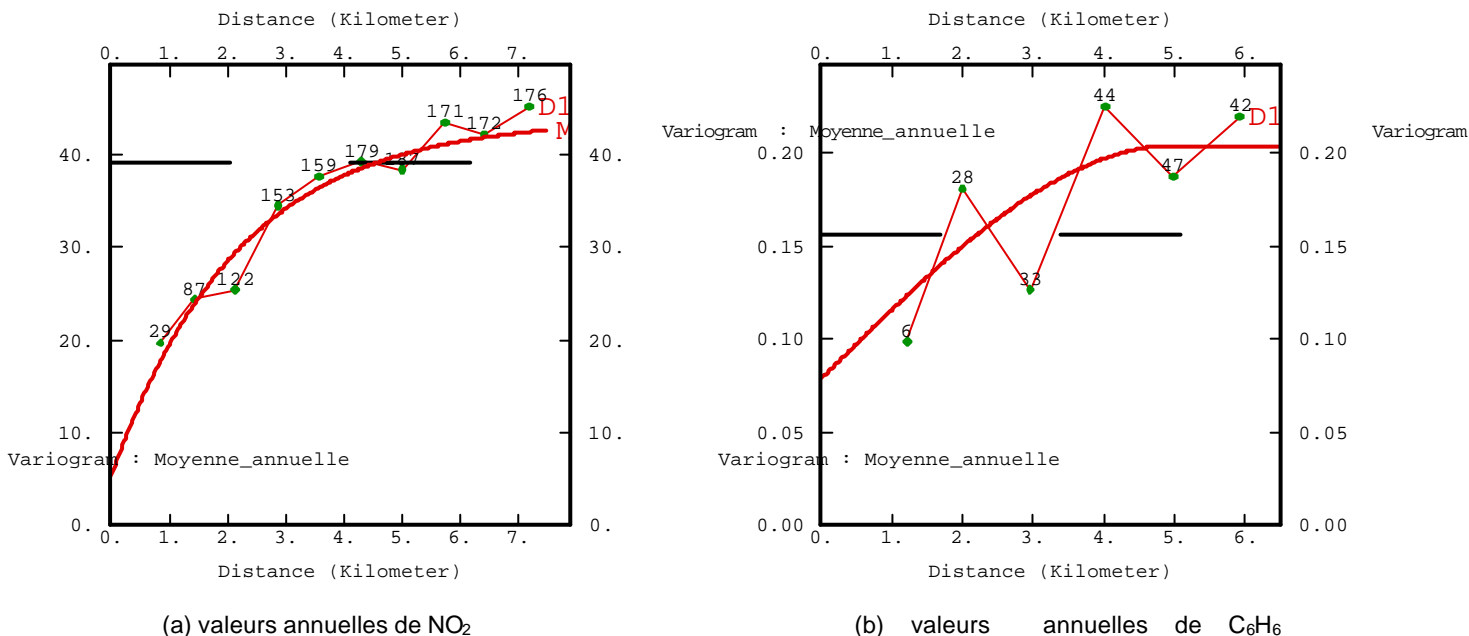
Seul l'algorithme d'interpolation géostatistique du krigeage intègre une règle de pondération directement déduite du comportement spatial de la donnée du polluant. Il permet en outre d'intégrer dans l'interpolation les effets de variables auxiliaires comme par exemple les émissions et/ou la densité de population. C'est cet algorithme qui été retenu ici pour réaliser les cartographies des champs de concentration en pollution sur l'aire géographique du SIVOM de l'agglomération mulhousienne.

# Cartographie des champs de concentration en dioxyde d'azote (NO<sub>2</sub>) et en benzène (C<sub>6</sub>H<sub>6</sub>) - Interpolation par krigeage ordinaire monovariante

## Analyse du comportement spatial des données et modélisation statistique

Dans le cadre de la méthode géostatistique du krigeage, la prise en compte des caractéristiques spatiales naturelles du polluant se fait à l'aide d'une analyse variographique encore dénommée analyse structurale. Cette analyse est fondamentale pour le choix du modèle d'interpolation à appliquer aux données expérimentales et elle n'est disponible que sous certains logiciels dédiés à la géostatistique comme ISATIS de la société GEOVARIANCES ou GEOSTATISTICAL ANALYST de la société ESRI.

L'analyse variographique permet de quantifier la variabilité spatiale du polluant à partir des valeurs observées aux différents sites ponctuels de mesure et en fonction de la distance qui les sépare. Le variogramme illustre ainsi la continuité ou corrélation spatiale du phénomène étudié si on l'inverse graphiquement pour obtenir une fonction décroissante (covariance). En pratique, plus la distance entre deux sites de mesure est grande et moins les mesures sont corrélées.



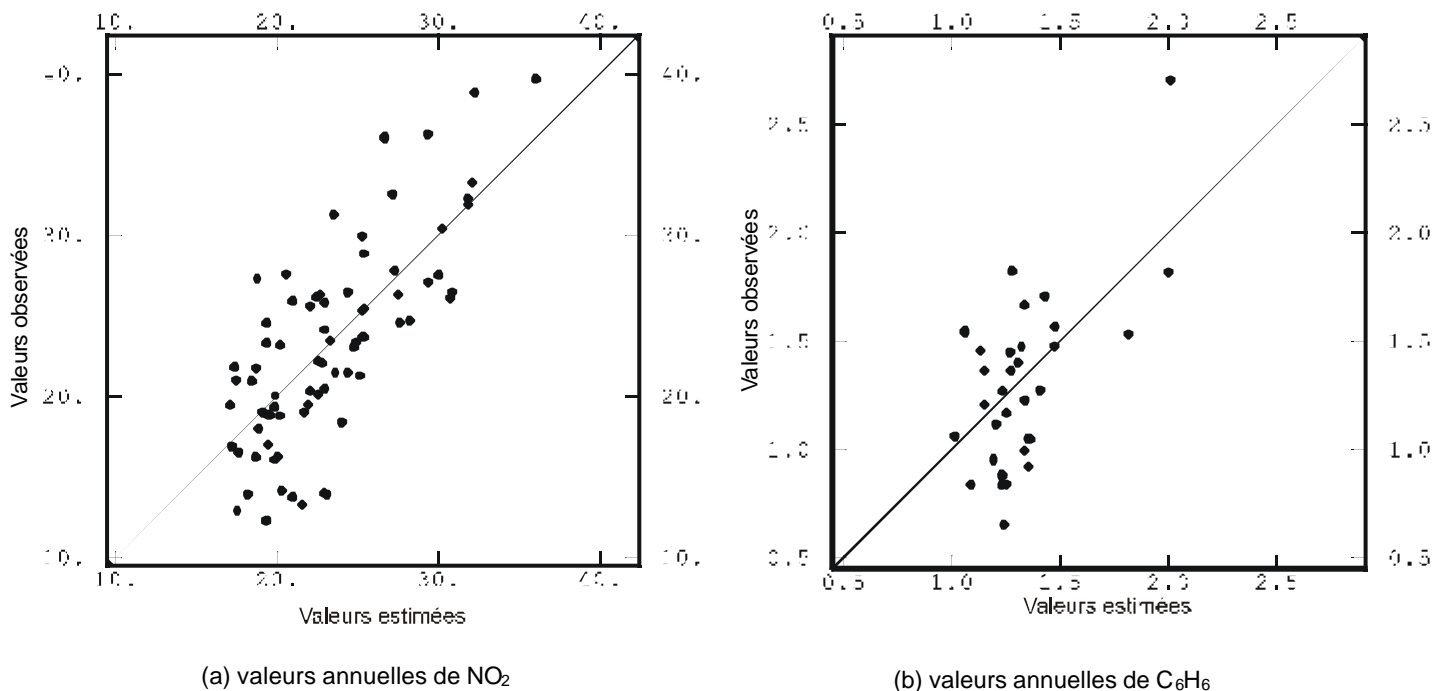
**Fig. 1** – Variogrammes expérimentaux (ligne mince) et modèles de variogramme (ligne épaisse) ajustés aux données de NO<sub>2</sub> (a) et de C<sub>6</sub>H<sub>6</sub> (b)

De ce fait, les variogrammes définis sur les valeurs de concentrations annuelles en NO<sub>2</sub> et C<sub>6</sub>H<sub>6</sub> mesurées sur l'ensemble des communes du SIVOM de Mulhouse font ressortir une meilleure corrélation dans l'espace du premier polluant comparativement au second (fig. 1a et 1b). Le polluant est non corrélé au-delà de la distance où le modèle de variogramme plafonne autour d'un palier atteint à environ 6-7 kilomètres pour le NO<sub>2</sub> et seulement 3,5-4,5 kilomètres pour le C<sub>6</sub>H<sub>6</sub>. Pour chacun de ces polluants, il n'existe donc plus de relation significative entre deux mesures au-delà de ces distances respectives. Le modèle de variogramme propre à l'un et l'autre des polluants entré dans un système linéaire d'équations (le système de krigeage) conduit à déterminer le poids optimal à affecter aux données pour l'interpolation. De cette manière, les données qui sont plus corrélées avec la concentration au nœud considéré reçoivent un poids plus grand que celles qui sont moins corrélées avec la concentration en ce même nœud.

Selon cette méthode, ce n'est plus la distance entre les sites de mesure et le nœud qui sont pris en compte, mais bien la corrélation spatiale du polluant entre les sites échantillonnés et la concentration au nœud considérée. Les informations déduites du modèle de variogramme établi sur la base de l'échantillon de mesures font que la méthode d'interpolation du krigeage s'adapte correctement au comportement spatial des données de pollution. La valeur estimée en un nœud correspond à la valeur la plus probable calculée à partir du modèle statistique calé et optimisé par rapport aux données expérimentales.

### Contrôle de la précision du modèle d'interpolation

Cette opération permet de vérifier l'adéquation entre les données et le modèle variographique utilisé. Elle consiste à estimer chaque donnée dont on connaît la valeur en l'éliminant des données d'entrée. La comparaison de la valeur estimée à la valeur observée ou valeur vraie se fait via un nuage de points dont la configuration autour de la bissectrice atteste la robustesse du modèle : plus le nuage de points est resserré autour de la bissectrice, meilleure est l'estimation des valeurs par krigeage.



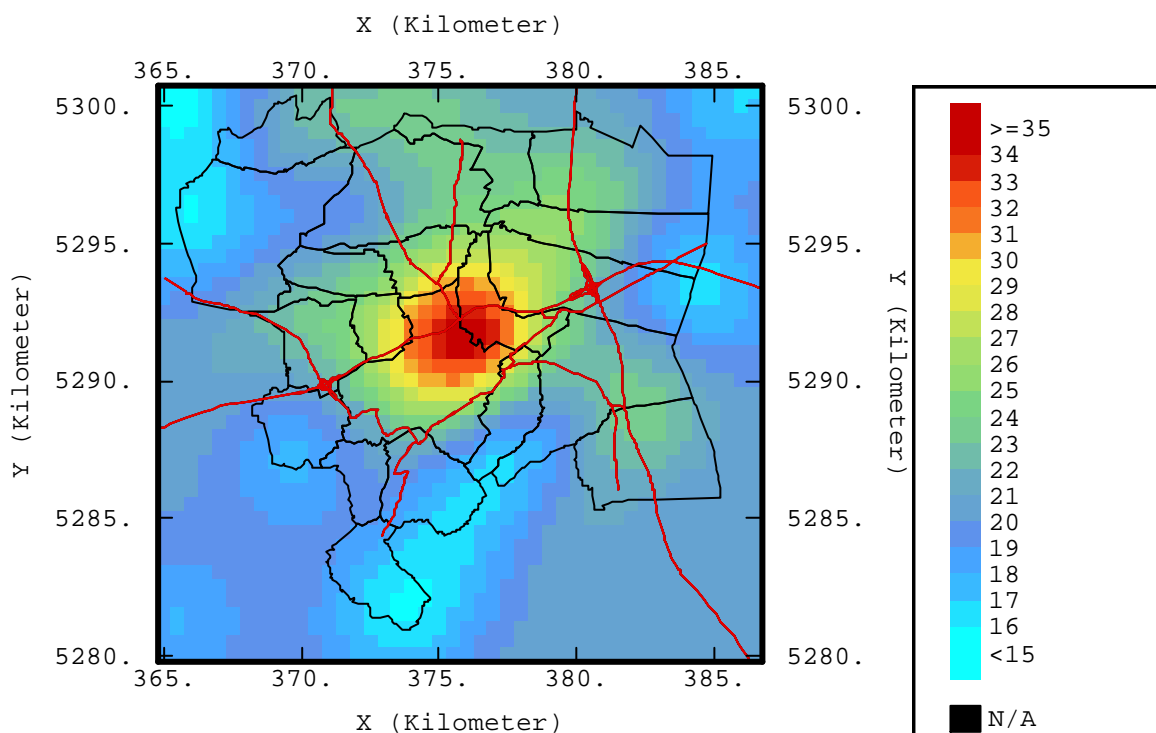
**Fig. 2** – Résultats de la validation croisée selon le modèle d'interpolation relatif au krigeage ordinaire monovariante pour les données de NO<sub>2</sub> (a) et de C<sub>6</sub>H<sub>6</sub> (b)

Les résultats de la validation croisée montrent que les modèles de variogramme construits pour évaluer et cartographier par krigeage les champs de concentration en NO<sub>2</sub> et C<sub>6</sub>H<sub>6</sub> sont relativement satisfaisants (fig. 2a et 2b). Soulignons toutefois que le modèle de variogramme établi pour les données du NO<sub>2</sub> va donner des résultats un peu plus lissés que ceux issus du modèle de variogramme appliqué aux données du C<sub>6</sub>H<sub>6</sub>. Cela est dû pour partie au fait que l'« effet de pépité » utilisé comme première structure du modèle de variogramme pour ajuster le comportement spécifique du polluant à l'origine tend à lisser d'avantage les valeurs extrêmes de concentration mesurée pour le NO<sub>2</sub> que pour le C<sub>6</sub>H<sub>6</sub>.

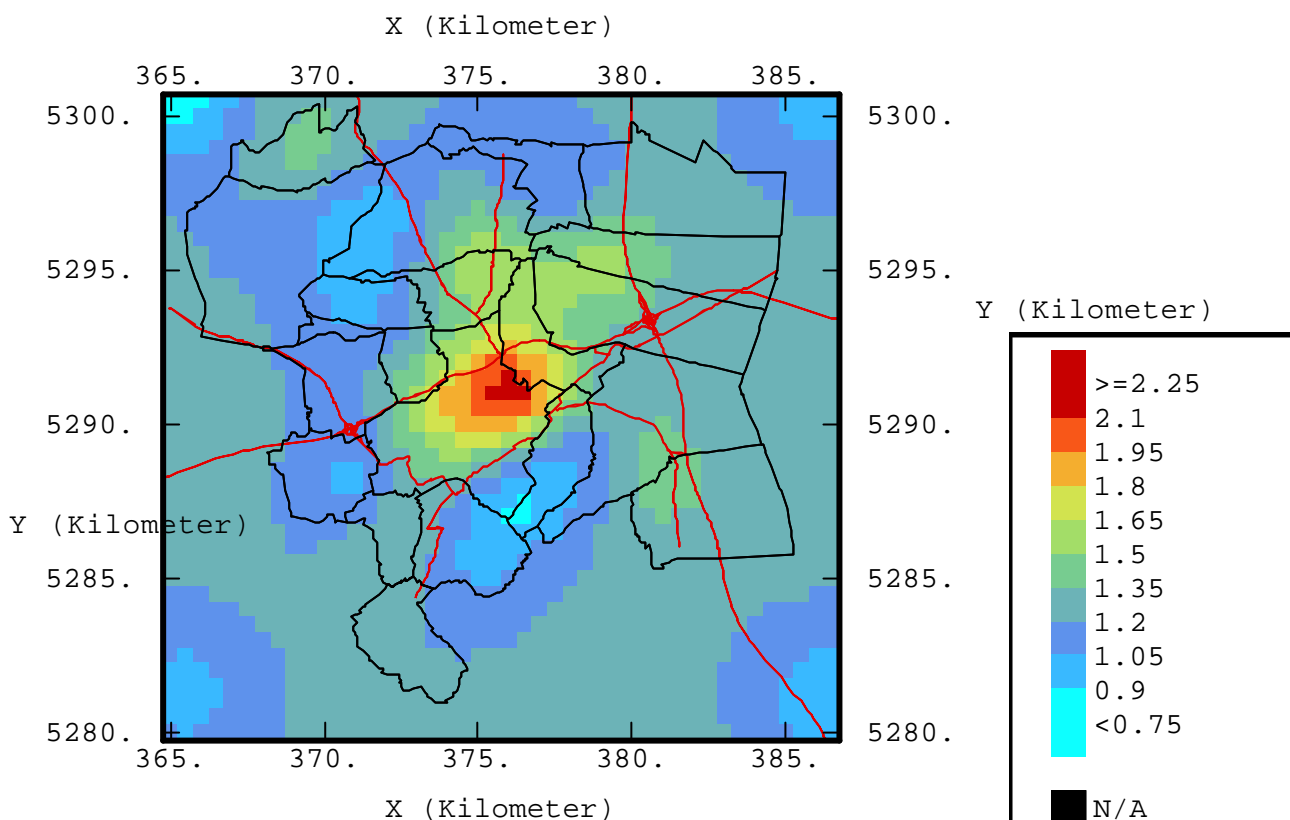
## Résultats cartographiques

Les résultats cartographiques sur les données de NO<sub>2</sub> et du C<sub>6</sub>H<sub>6</sub> obtenus à l'aide du krigeage ordinaire monovariante sont présentés ci-dessous (cartes 1 et 2). L'échelle des valeurs correspond à la moyenne annuelle des concentrations en µg/m<sup>3</sup> calculée à partir des deux phases de campagnes de mesures estivales et hivernales réalisées par tubes passifs.

Les champs de pollution de fond ainsi cartographiés reflètent bien une certaine «logique terrain», tant dans la répartition spatiale du NO<sub>2</sub> que du C<sub>6</sub>H<sub>6</sub>. Les concentrations les plus fortes se retrouvent localisées sur l'hyper-centre de l'agglomération mulhousienne et vont décroissantes à mesure que l'on s'en éloigne. Il est à noter que les gradients de pollution se différencient selon la situation au vent ou sous le vent du panache polluant de la ville de Mulhouse. Il ressort notamment un resserrement du gradient de pollution du côté sud-ouest de l'agglomération comparativement au côté nord-est. Cette configuration est liée de surcroît aux vents dominants qui durant la campagne ont soufflé selon cet axe directionnel privilégié par le contexte topographique de la vallée du Rhin. On observe par ailleurs des couloirs de pollution qui s'étendent en direction de principaux secteurs urbanisés et grandes voies de circulation (A36, D366). On remarquera également une décroissance plus rapide du gradient de pollution pour le C<sub>6</sub>H<sub>6</sub> que pour le NO<sub>2</sub> à mesure que l'on s'éloigne de l'hyper centre de l'agglomération qui confirme les schémas spécifiques de corrélation spatiale établis par l'analyse variographique sur ces deux polluants.



**Carte 1** – Répartition spatiale de la pollution de fond en NO<sub>2</sub> sur les communes du SIVOM de Mulhouse – Krigeage ordinaire monovariante - Moyenne annuelle en µg/m<sup>3</sup>

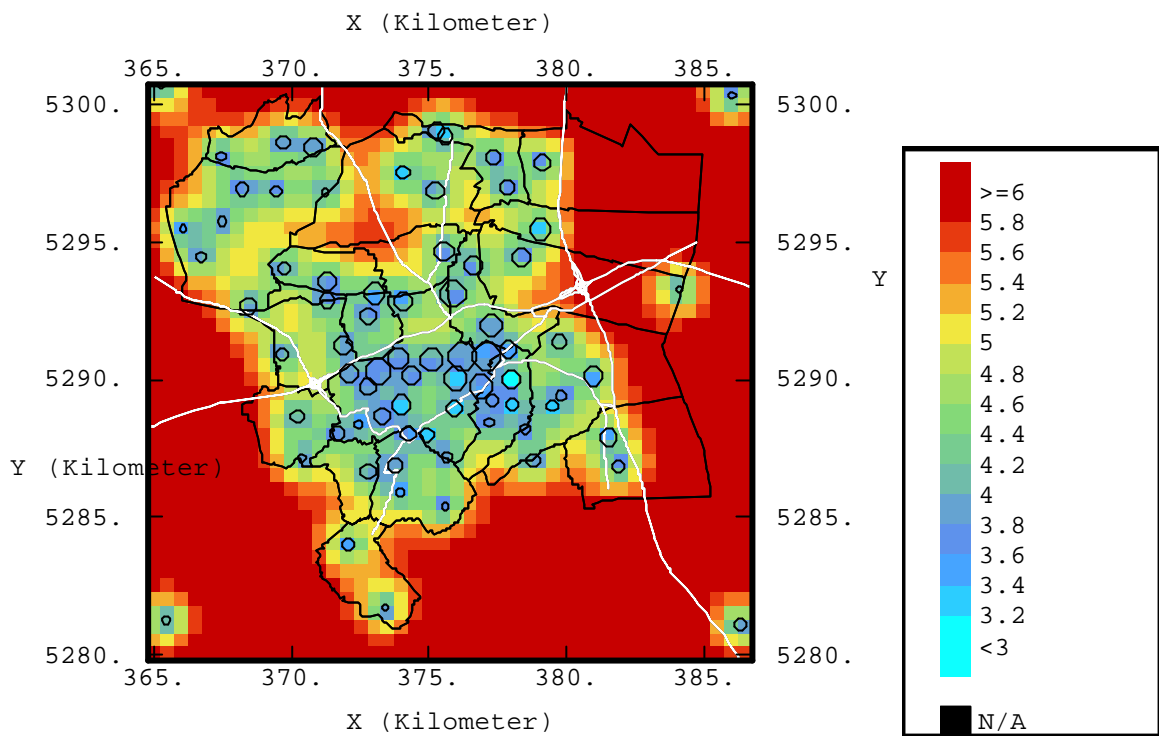


**Carte 2** – Répartition spatiale de la pollution de fond en  $\text{C}_6\text{H}_6$  sur les communes du SIVOM de Mulhouse – Krigeage ordinaire monovariante - Moyenne annuelle en  $\mu\text{g}/\text{m}^3$

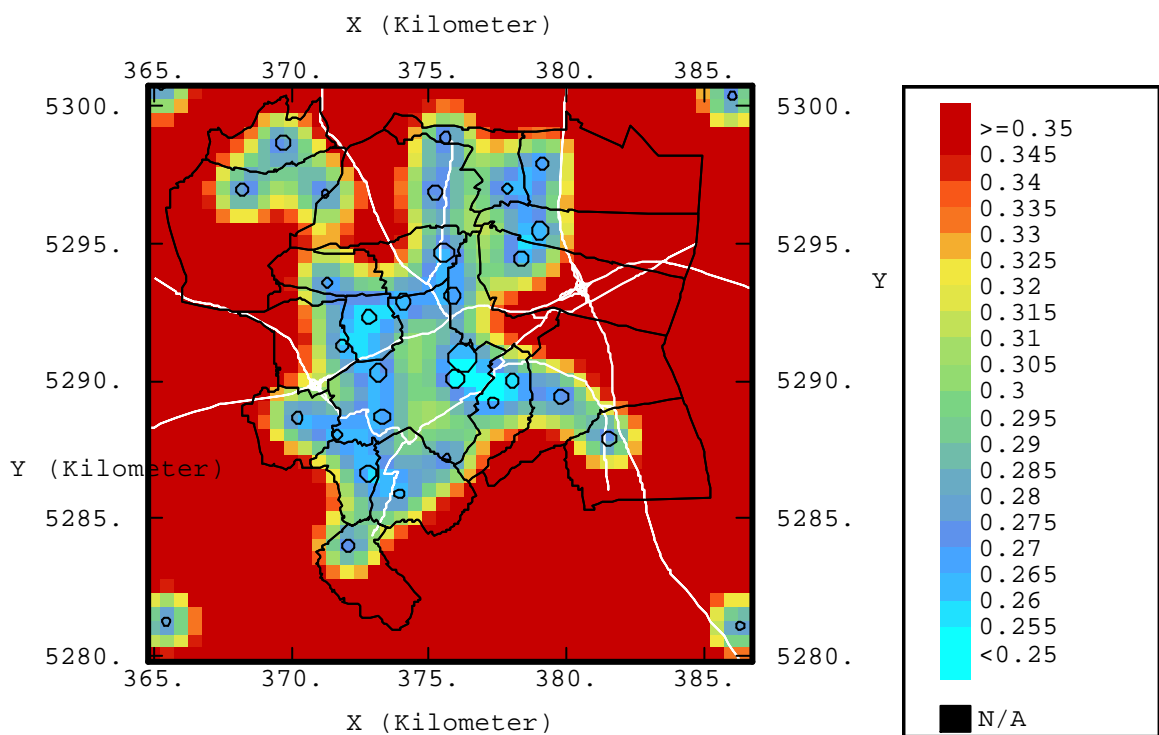
Les cartes d'écart-types de krigeage permettent de quantifier a posteriori l'incertitude associée à chaque carte en plus de l'analyse de validation croisée présentée ci-dessus. Ces cartes sont construites à la base à l'aide de la variance de l'erreur calculée à partir de la méthode probabiliste d'interpolation du krigeage. Les résultats exprimés en terme de racine carré de la variance, soit l'écart-type de krigeage, caractérisent la variabilité ou dispersion possible de l'erreur d'interpolation dans l'unité même de la variable cartographiée.

Ainsi, un écart-type de krigeage élevé indiquera une valeur interpolée incertaine. A l'inverse, un écart-type de krigeage faible indiquera une valeur interpolée en moyenne proche des observations et une bonne précision de la cartographie. Soulignons que la valeur de l'écart-type de krigeage tend à croître très rapidement à distance d'un point renseigné par une mesure car cette variable ne dépend que de la forme du variogramme et de la géométrie des données.

Les cartes d'écart-types de krigeage sont donc à interpréter avec beaucoup de précautions : elles sont très sensibles au type du modèle de variogramme retenu pour l'interpolation (surtout lorsque ce dernier comporte une structure « pépétique » à l'origine comme dans le cas présent). La carte d'écart-type de krigeage du  $\text{NO}_2$  est ainsi beaucoup plus empreinte par « l'effet pépétique » introduit dans le modèle de variogramme que celle associée au  $\text{C}_6\text{H}_6$  (cartes 3 et 4). Il importe par ailleurs de préciser que les cartes d'écart-type de krigeage ne sont pas toujours directement comparables selon que l'interpolation a été établie sur des informations ponctuelles ou sur des informations moyennées sur une surface donnée (krigeage ponctuel et krigeage de blocs).



**Carte 3** – Ecart-type de krigeage sur les données annuelles en  $\text{NO}_2$  – Krigeage ordinaire monovariante



**Carte 4** – Ecart-type de krigeage sur les données annuelles en  $\text{C}_6\text{H}_6$  – Krigeage ordinaire monovariante

## Intégration d'informations indirectes liées à des variables auxiliaires

L'ASPA dispose sous SIG d'un certain nombre de plans d'information spatialisée relatifs en particulier aux émissions de polluants et aux données du recensement général de la population en Alsace réalisé en 1999 par l'INSEE. Ces données sont censées présenter une certaine relation avec la concentration des polluants considérés. La prise en compte d'informations indirectes associées aux variables auxiliaires a pour but d'améliorer la représentation cartographique notamment dans les zones pauvres, voire dépourvues en mesure. L'intégration de l'information liée à la variable auxiliaire dans le système géostatistique d'interpolation permet en quelque sorte de compenser le sous-échantillonnage et/ou l'absence de mesure du polluant en certains lieux.

### Cartographie des champs de concentration en dioxyde d'azote (NO<sub>2</sub>) et en benzène (C<sub>6</sub>H<sub>6</sub>) avec prise en compte de données auxiliaires - Interpolation par krigeage multivariable

#### *Choix de la méthode d'interpolation en krigeage multivariable*

On dispose de deux techniques pour intégrer des variables auxiliaires dans la méthode d'interpolation : la technique de co-krigeage et la technique de krigeage avec dérive externe.

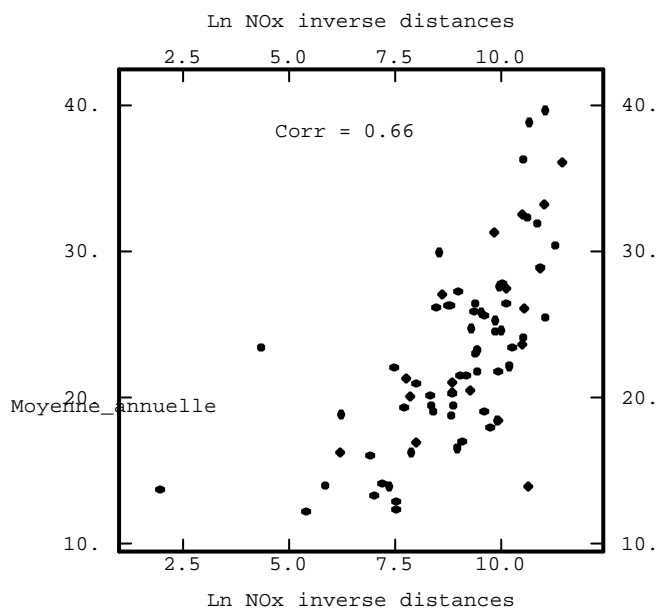
Dans le cas du co-krigeage, il est nécessaire d'élaborer un modèle structural ou variographique multivariable décrivant à la fois la corrélation spatiale de chacune des variables considérées séparément (variable d'intérêt : le polluant – variable auxiliaire : ex. les émissions) et leur corrélation conjointe encore dénommée corégionalisation. Cela suppose l'existence d'un lien direct et linéaire de type positif entre les deux variables. La mise en œuvre du co-krigeage implique ainsi de modéliser les variogrammes simples et croisés des variables d'intérêt et auxiliaire afin de les entrer dans le système de co-krigeage.

Dans le cas du krigeage avec dérive externe, on se place dans l'hypothèse où la variable d'intérêt présente une structure d'ensemble modelée par une autre variable. En d'autres termes, on considère que l'on dispose d'une fonction de forme, c'est-à-dire d'une variable régionalisée distincte de la variable d'intérêt, capable de décrire les grandes lignes structurales de la variable à estimer. Le comportement spatial de la variable auxiliaire s'apparente alors à un indicateur de la tendance générale, dite dérive externe, représentative des répartitions des concentrations en pollution sur le domaine géographique considéré.

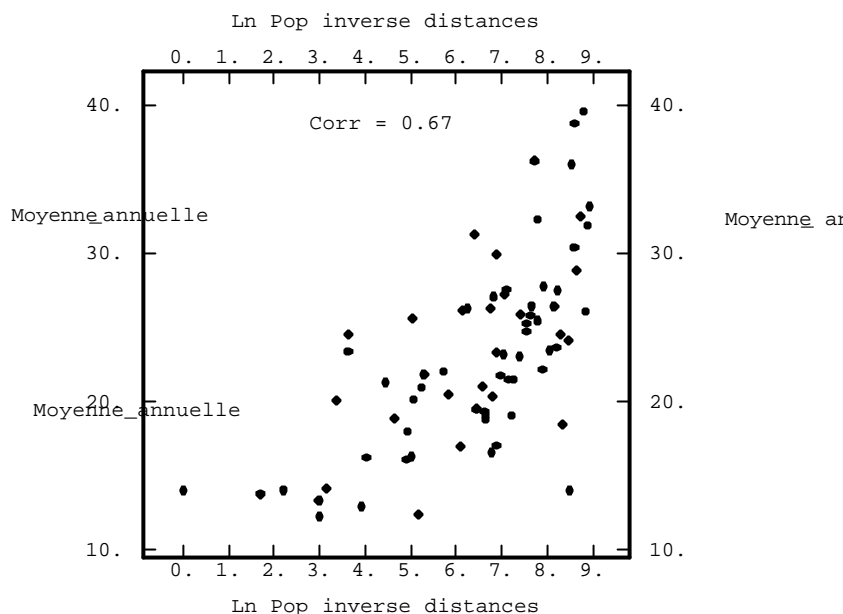
L'analyse des relations entre les concentrations annuelles des polluants et respectivement les émissions et la densité de population calculées sur une maille de 500 m<sup>2</sup> (correspondant à la maille de reconstitution des champs de concentration de pollution en NO<sub>2</sub> et C<sub>6</sub>H<sub>6</sub>) est illustrée en figures 3a à 3d. Elle a été menée sur la base d'une « transformée » des variables auxiliaires de type : Log népérien (val var aux + 1). Cette transformation a été réalisée dans le but de ramener l'ordre de grandeur des unités de mesures des variables auxiliaires sur une base plus proche de celle des variables d'intérêt.

Les nuages de points caractérisent l'existence d'une corrélation plus apparentée à une dérive car de structure faiblement linéaire. Dans ce cas de figure correspondant à un modèle non stationnaire, la technique du krigeage avec dérive externe est la plus adaptée pour appliquer la méthode d'interpolation multivariable. C'est cette méthode qui a été retenue pour produire les cartes finales de la répartition des champs de pollution de fond en NO<sub>2</sub> et C<sub>6</sub>H<sub>6</sub> sur l'ensemble des communes du SIVOM de Mulhouse (cf. p.n dans corps du rapport).

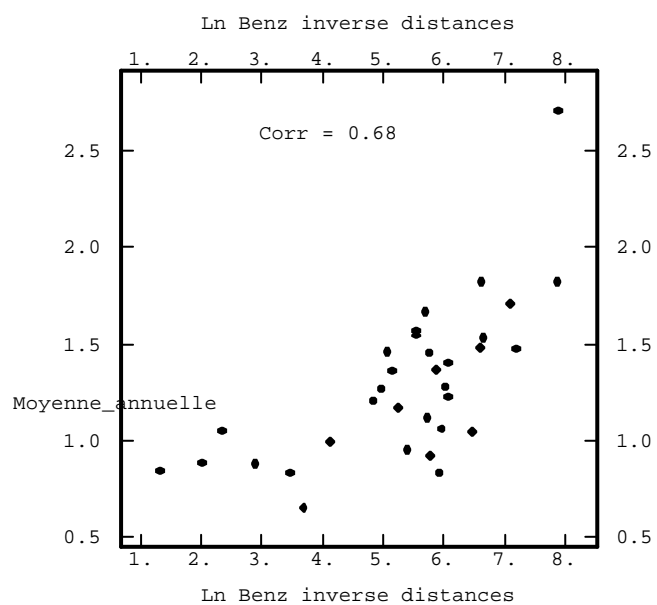
Après modélisation de la forme de la dérive externe associée aux variables auxiliaires des émissions en NO<sub>x</sub> et en benzène d'une part, et à la densité de la population d'autre part, on introduit ses composantes dans le système de krigeage. Plusieurs modèles sont ainsi testés et classés par ordre décroissant de performance. Les poids affectés aux données de polluant lors de l'interpolation sont calculés de telle sorte à respecter au mieux l'allure de la dérive sur le champ et à coïncider à la meilleure covariance généralisée.



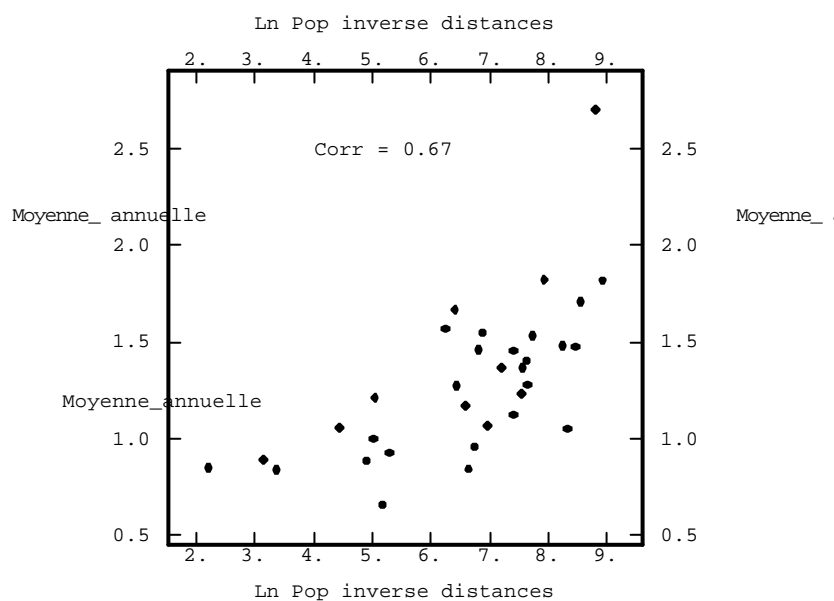
(a) valeurs annuelles de NO<sub>2</sub> et Ln NOx



(b) valeurs annuelles de NO<sub>2</sub> et Ln Pop 99



(c) valeurs annuelles de C<sub>6</sub>H<sub>6</sub> et Ln Benz



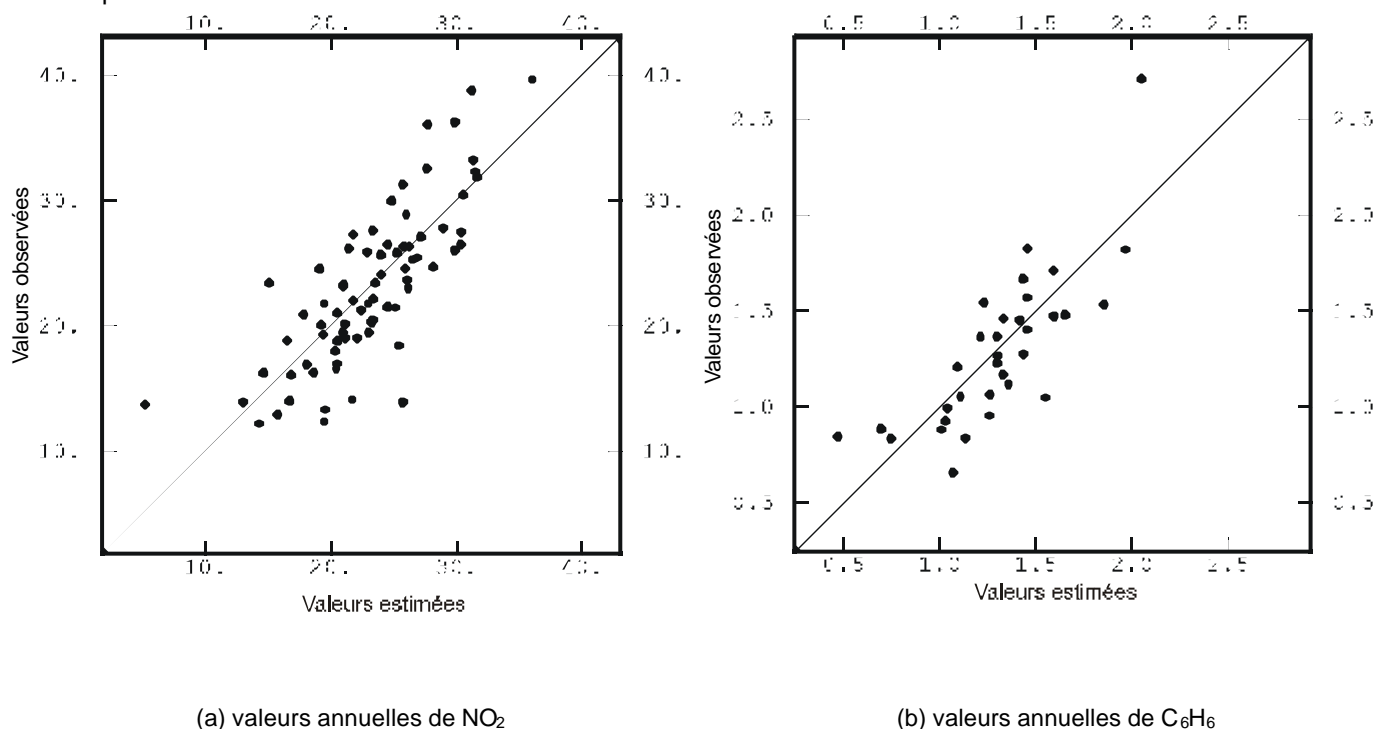
(d) valeurs annuelles de C<sub>6</sub>H<sub>6</sub> et Ln Pop 99

**Fig. 3** – Analyse des relations entre les variables auxiliaires des émissions et de la densité de population et les données de NO<sub>2</sub> (a et b) et de C<sub>6</sub>H<sub>6</sub> (c et d)



### Vérification de la précision et de l'optimisation du modèle d'interpolation

L'évaluation de la pertinence du modèle retenu in fine se fait comme précédemment par l'examen des résultats de la validation croisée. Ceux-ci, présentés en figures 4a et 4b, font clairement ressortir que la prise en compte de la dérive externe afférente aux informations indirectes des émissions et de la densité de la population améliore sensiblement la précision des estimations des valeurs de concentrations. Les nuages de points valeurs vraies/estimées pour chacun des polluants apparaissent beaucoup plus resserrés autour de la bissectrice que dans le cas des modèles de krigeage monovarié présentés ci-dessus (cf. fig. 2a et 2b). Cela atteste l'optimisation de la méthode d'interpolation pour une représentation cartographique plus réaliste des champs de concentration des polluants considérés.

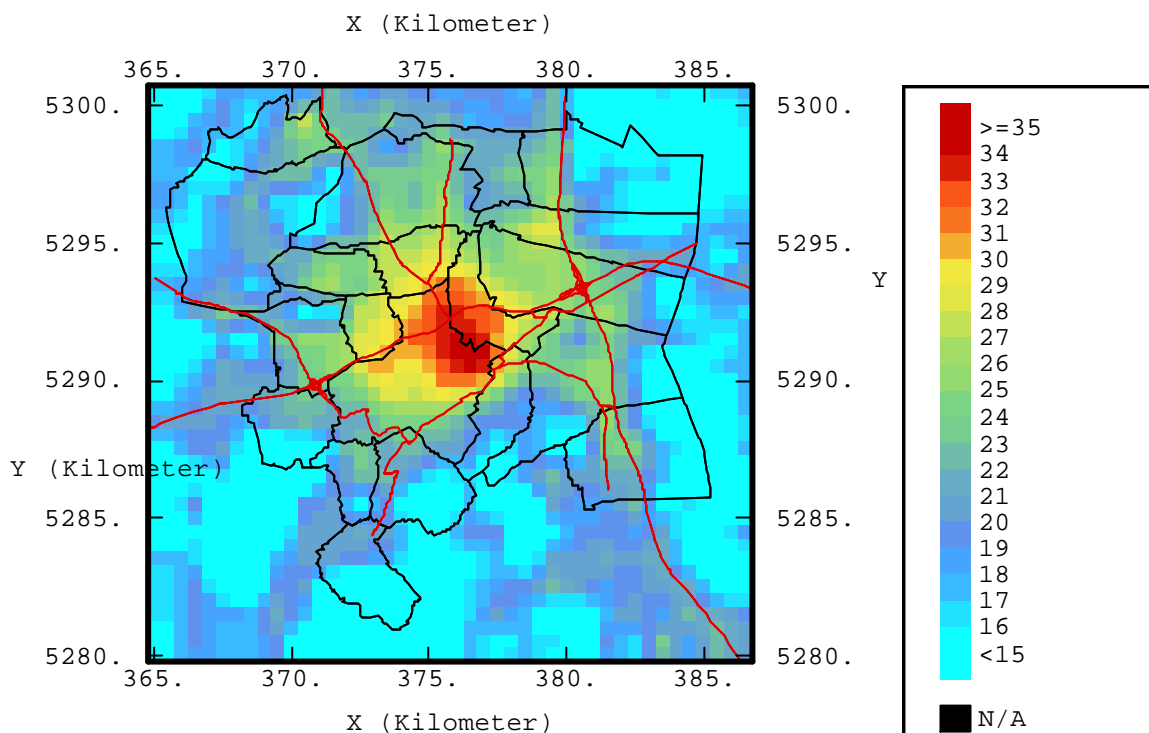


**Fig. 4** – Résultats de la validation croisée selon le modèle d'interpolation relatif au krigeage multivarié avec dérive externe pour les données de NO<sub>2</sub> (a) et de C<sub>6</sub>H<sub>6</sub> (b)

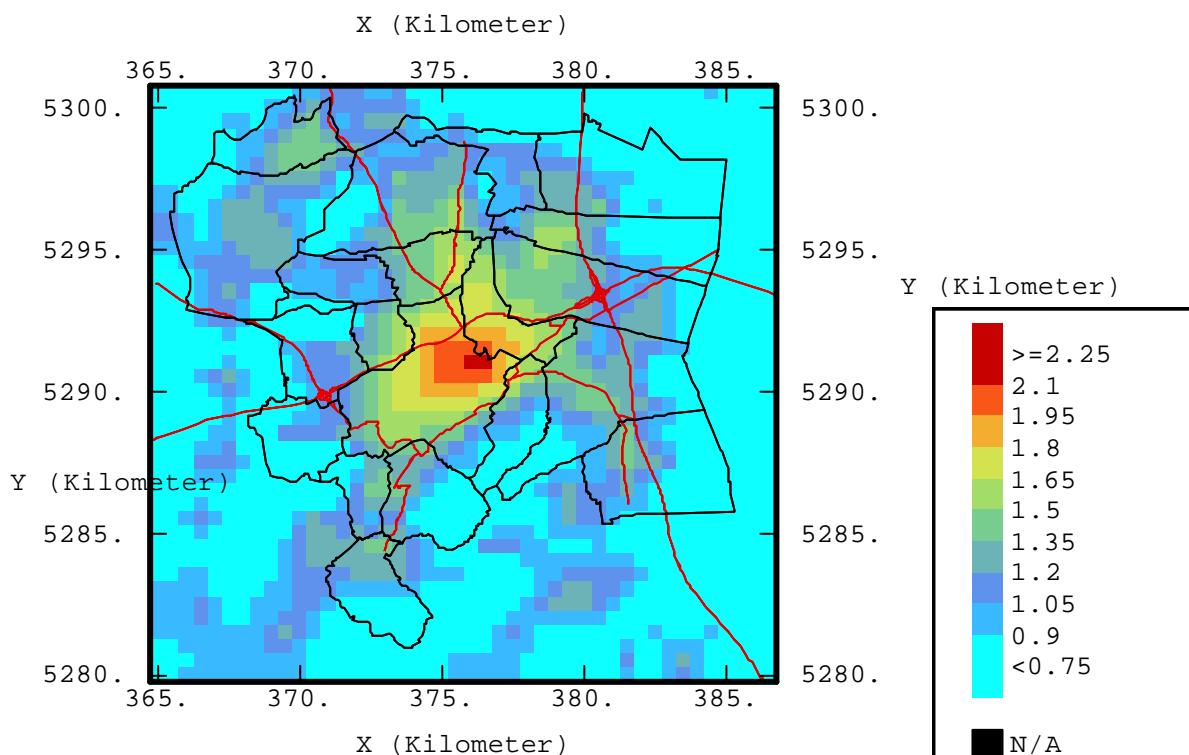
## Résultats cartographiques

Les cartes obtenus à l'aide du krigeage multivariable avec dérive externe sont présentées ci-dessous (cartes 5 et 6). Au vu des résultats, on constate que cette méthode d'interpolation débouche effectivement sur une représentation plus réaliste des champs de concentration en pollution de fond par le  $\text{NO}_2$  et le  $\text{C}_6\text{H}_6$  que celle issue du krigeage monovariable globalement trop lisse. Cela se remarque sur le périmètre central entre les sites de mesure, dans le secteur Nord ouest du domaine géographique moins densément instrumenté et dans les zones extrêmes sud-ouest, sud-est et nord-est totalement dépourvues de mesure où l'on travaille en « pure extrapolation ». On observe entre autres un réhaussement significatif des concentrations sur les secteurs urbanisés ainsi qu'au voisinage des infrastructures des voies de circulation automobile (en particulier pour le  $\text{NO}_2$  de caractère mixte). Dans les zones peu ou pas échantillonnées, la structure spatiale des concentrations reflète à des degrés divers celle des variables auxiliaires mise en trame de fond.

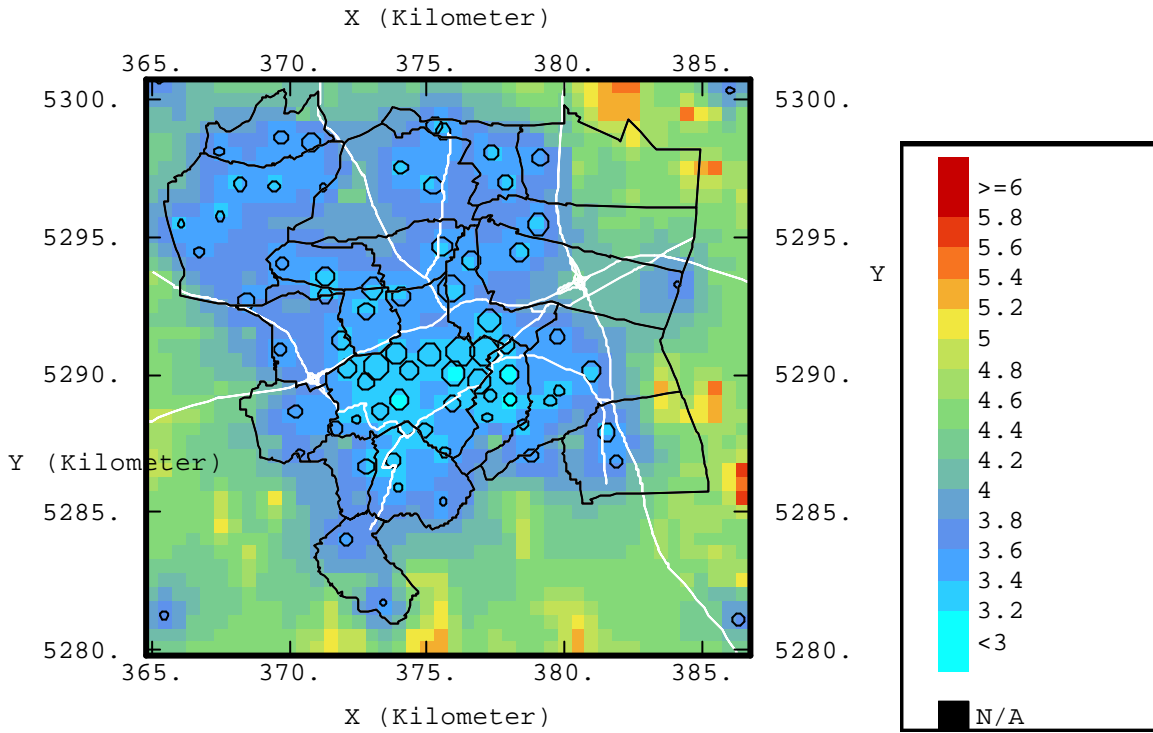
Les cartes de l'écart-type de krigeage confirment l'optimisation de la cartographie établie à l'aide du krigeage multivariable avec dérive externe (cartes 7 et 8). En l'occurrence, l'« effet pépitique » constaté sur la carte d'écart-type de krigeage monovariable est très fortement atténué. Dans tous les cas, la précision des cartes est grandement améliorée, tant dans les secteurs instrumentés que ceux non instrumentés. Soulignons qu'en krigeage avec dérive externe, l'écart-type dépend aussi de la forme de la dérive. De ce fait, moins la dérive s'assimile au comportement local des données de pollution, plus l'incertitude associée à la carte peut devenir importante. Il est à noter par ailleurs que la moyenne de la variance de l'erreur tend à croître proportionnellement avec la taille de l'échantillon ainsi qu'avec l'amplitude des valeurs extrêmes associées au jeu de données considéré (amplitude plus faible dans l'échelle des valeurs de concentrations  $\text{C}_6\text{H}_6$  à  $\text{NO}_2$ ).



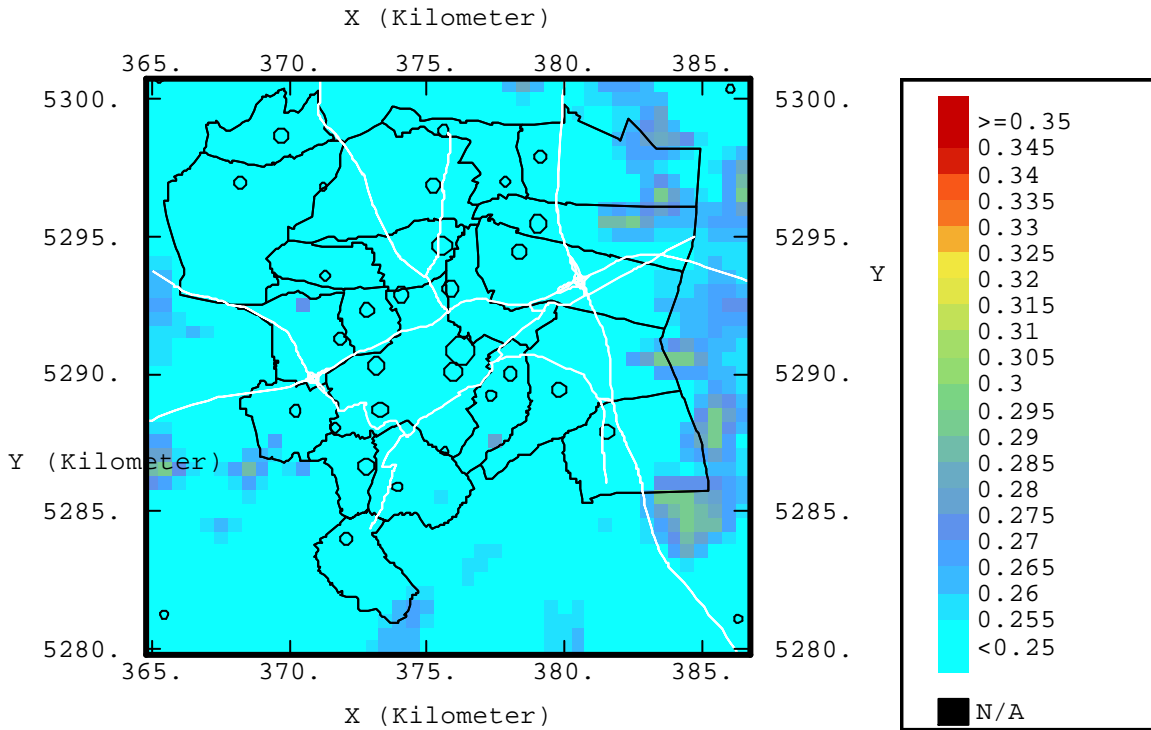
**Carte 5** – Répartition spatiale de la pollution de fond en  $\text{NO}_2$  sur les communes du SIVOM de Mulhouse – Krigeage multivariable avec dérive externe - Moyenne annuelle en  $\mu\text{g}/\text{m}^3$



**Carte 6** – Répartition spatiale de la pollution de fond en  $\text{C}_6\text{H}_6$  sur les communes du SIVOM de Mulhouse – Krigeage multivariable avec dérive externe - Moyenne annuelle en  $\mu\text{g}/\text{m}^3$



**Carte 7** – Ecart-type de krigeage sur les données annuelles en  $\text{NO}_2$  – Krigeage multivariable avec dérive externe



**Carte 8** – Ecart-type de krigeage sur les données annuelles en  $\text{C}_6\text{H}_6$  – Krigeage multivariable avec dérive externe

## Conclusion

La méthode d'interpolation géostatistique adoptée pour reconstituer les champs de concentration de pollution de fond en  $\text{NO}_2$  et  $\text{C}_6\text{H}_6$  sur l'ensemble des communes du SIVOM de l'agglomération mulhousienne a permis d'obtenir une cartographie réaliste et quantifiable quant à sa précision. La modélisation du comportement spatial des données observées et l'intégration de variables auxiliaires aident à respecter et à coller au mieux aux « aspects-terrain » du phénomène étudié. La quantification de l'incertitude associée à la carte répondent aussi bien aux besoins du système Assurance Qualité dans la production et le traitement des données qu'à celles du public et des décideurs.

L'intégration à terme sous SIG de tous les résultats cartographiques issus des campagnes de mesure réalisées sur l'unité urbaine de Mulhouse devrait permettre une connaissance actualisée des modalités d'exposition de la population résidente de ce secteur par rapport au dépassement de normes des polluants relatifs notamment au dioxyde d'azote et au benzène. Cette connaissance sera de plus affinée sous l'angle des fluctuations saisonnières grâce aux résultats recueillis aux différentes saisons (été/hiver).

Pour aborder les questions d'exposition de la population à la pollution de fond, il est envisagé d'utiliser des techniques adaptées telles celles correspondantes aux simulations conditionnelles. En outre, des travaux vont être engagés pour améliorer et automatiser certains traitements de la cartographie géostatistique. L'objectif consiste notamment à réduire l'information pouvant découler des divers plans d'information spatialisée disponibles à l'ASPA (modèle numérique de terrain, plan d'occupation des sols, cadastre des émissions, ....) et à utiliser un test statistique pertinent pour la sélection des variables auxiliaires à prendre en compte dans la cartographie des différents polluants.